

Enquêtes par questionnaire : les valuateurs de représentations

Alain DUBUS

Université Charles-de-Gaulle — Lille III

Équipe THÉODILE – E.A.1764

1 CONTEXTE

Le contexte étroit des réflexions exposées ci-dessous est celui de l'enseignement de la méthodologie de l'enquête par questionnaires en Maîtrise de Sciences de l'Éducation, et du conseil aux étudiants de 3^{ème} cycle. La maîtrise, avec le poids central de la réalisation du mémoire, suppose simultanément l'apprentissage des techniques de recherche et leur mise en œuvre immédiate, et pose de redoutables problèmes de priorités circulaires entre choix du sujet, question de recherche, cadre théorique, problématique et hypothèses, recueil de données, saisie et exploitation statistique, interprétation et vérification des hypothèses ... et retour à la case départ. A chacune des étapes qui viennent d'être citées, une connaissance insuffisante des problèmes spécifiques aux étapes *ultérieures* est de nature à détériorer gravement la qualité du résultat final. Que faire ? Le champ méthodologique est foisonnant, et subordonner la mise en route d'un premier projet de recherche à une appropriation cohérente et autonome des techniques, tant dans leur utilisation que dans le choix parmi celles qui s'offrent, revient à décourager l'entreprise, d'autant que l'énergie nécessaire à surmonter certaines difficultés d'approche des aspects techniques ne trouve souvent sa source que dans l'intérêt de la question étudiée. Toute aussi déprimante est la solution consistant à « faire semblant », c'est-à-dire à formuler sur tous les aspects de la tâche des exigences de bon niveau, quitte à valider *in fine* des réalisations plutôt dégradées au regard de ces exigences. Une solution intermédiaire pourrait consister à proposer à une première appropriation un éventail réduit de techniques robustes, de manière à en rendre possible l'utilisation heureuse dans un délai raisonnable, en reportant à d'éventuelles phases ultérieures de la formation à la recherche les diversifications méthodologiques souhaitables.

Le contexte plus large des mêmes réflexions est l'existence d'une pratique assez intensive du questionnaire dans les différentes sphères du travail social, éducatif ou culturel. A tort ou à raison, le questionnaire s'impose souvent spontanément comme

l'instrument le plus adapté à l'acquisition d'informations concernant les publics, à l'évaluation d'actions, voire à l'extraction de nouveaux besoins. Des trésors d'énergie et d'imagination sont fréquemment dépensés par des acteurs de terrain, avec des résultats parfois très décevants : exploitation indigente, obstacles insurmontables au codage liés à une formulation défectueuse des questions ... Il ne s'agit pas ici de faire le procès de ces tentatives parfaitement respectables et qui méritent parfois d'être valorisées, mais de déplorer un gaspillage d'efforts. Une vision stéréotypée des étapes du travail, elle-même sous-tendue par une représentation implicite de la division technique et sociale du travail de recherche, amène parfois les équipes qui se sont placées dans la situation peu enviable d'avoir recueilli des données et de ne savoir qu'en faire à se retourner vers des spécialistes universitaires de la méthodologie, pour s'entendre expliquer avec ménagements qu'il est bien trop tard pour rectifier quoi que ce soit. A vrai dire, il peut même arriver à des thésards d'éprouver de semblables déconvenues.

La lutte contre ce genre de choses est une de mes occupations constantes depuis de nombreuses années déjà. Il y a, en effet, matière à lutter. Contre la division technique et sociale du travail de recherche, contre l'étanchéité des étapes et des *moments* de la recherche. Pour l'autonomie instrumentale du chercheur, qui suppose d'avoir en vue, en même temps que l'objet et le cadre théorique dans lequel on entend le traiter, les techniques de recueil de données, les perspectives de traitement et les limites qu'elles imposent à l'interprétation. Ce qui suit est, sur un objet particulier et restreint, une illustration de cette démarche.

2 UN PROBLÈME RÉCURRENT

2.1 LE CHAMP DES REPRÉSENTATIONS

La notion même de questionnaire recouvre des réalités assez variées. Il est des questionnaires exploratoires, assez proches d'une grille d'entretien directif, riches en questions ouvertes, prélude à une éventuelle investigation plus systématique, mais sans que l'exploitation quantitative soit la seule ni même la première préoccupation. Ce n'est pas de ce genre de questionnaires qu'il est question ici, mais de ceux qui sont explicitement construits pour recueillir le bénéfice du nombre, c'est-à-dire ceux pour lesquels on espère que le caractère immanquablement réducteur des questions fermées sera compensé par les potentialités statistiques d'un effectif important de répondants.

A l'intérieur même de cette catégorie de questionnaires résolument quantitatifs en termes d'effectifs visés, il ne s'agit pas non plus de n'importe quelle sorte de question. Les questions visant à recueillir des éléments factuels présentent leurs propres difficultés, qui ne concernent pas le présent propos. On se borne ici au champ des *représentations*, méta-catégorie neutre susceptible de rassembler tout ce qu'on pour-

rait énumérer avec les vocables d'opinions, d'attitudes, de jugements, d'évaluations, de positionnements, d'expressions idéologiques, de degrés d'adhésion ... qui ont en commun de solliciter la réaction subjective de l'enquêté à un énoncé.

Brutalement, on essaie de savoir ce que les gens pensent, mais on ne le saura qu'à travers ce qu'ils veulent bien dire, à supposer même qu'ils aient bien compris dans la question ce que l'enquêteur a cru y écrire. En d'autres termes, on sait d'avance que cette classe de questionnement va être truffée de pièges linguistiques, de problèmes de présentation de soi, et de ce qu'on pourrait appeler la négociation du rapport de force entre l'enquêteur et l'enquêté, ou encore la construction de l'enquêté, au sens où celui-ci n'est pas une entité préalablement disponible, mais l'un des pôles d'une relation à construire. En effet, pour qu'un questionné réponde à un questionneur, il faut qu'il y ait intérêt, qu'il en ait envie parce que le thème du questionnement résonne avec ses préoccupations, ou au minimum qu'il éprouve une certaine neutralité bienveillante envers l'investigateur. C'est que la relation est fondamentalement dissymétrique, et, à bien des égards, intrusive. La sanction usuelle du questionnaire diffusé dans des contextes où aucun de ces éléments favorisant n'est présent est un taux ridicule de réponses, éventuellement encore biaisé par le fait que ceux qui répondent à un questionnaire dans ces conditions peuvent être soupçonnés de n'être pas n'importe qui, mais d'appartenir à la catégorie de ceux qui écrivent facilement aux journaux. Dans le cas même, fréquent dans le domaine éducatif, où l'on enquête auprès d'un public captif, c'est-à-dire institutionnellement privé de la possibilité de ne pas répondre, les capacités de l'enquêté à s'évader de l'emprise de l'enquêteur ne sont pas nulles, et une partie de l'art de l'enquêteur réside dans la connaissance des ruses et contre-ruses qu'implique sa démarche.

2.2 LES STRATÉGIES SPONTANÉES DE QUESTIONNEMENT

Parmi les étudiants de Sciences de l'Éducation, comme parmi les praticiens évoqués plus haut (ce sont parfois les mêmes), des notions immédiates de ce que doit être un questionnaire sont souvent disponibles, nourries par de premières initiations aux méthodes d'enquête, mais aussi par les modèles des sondages omniprésents et des formulaires administratifs. En matière de représentations, ces notions inspirent parfois des postures de questionnement aussi naïves que brutales, mais aussi des raffinements qui, efficaces ou non, témoignent au moins de la prise de conscience que les choses ne vont pas de soi.

2.2.1 AU-DELÀ DU BRUTAL OUI-NON

Le modèle le plus élémentaire est celui de la question à laquelle on est invité à répondre par *oui* ou par *non*. Le souhait d'obtenir ainsi des positions aussi clairement tranchées que possible se heurte souvent au fait que l'opinion de l'enquêté le conduirait à répondre « ça dépend ». De là un nombre important de non-réponses, ou

de réponses « ne sait pas », si on a eu la bonté de les prévoir. Pire, peut-être : parmi les réponses *oui* et *non*, on peut suspecter que figurent les cas d'enquêtés d'humeur docile qui acceptent de « jouer le jeu » coûte que coûte en basculant d'un côté ou de l'autre, même si leur position est indécise. Il n'existe évidemment aucun moyen de discerner un *oui* de complaisance d'un vrai *oui* : ainsi la stratégie visant à « forcer » le répondant échoue dans le moment même où elle croit réussir, parce que l'interprétation des résultats obtenus est irréductiblement sujette à caution.

Le sentiment de l'inutile violence de ce modèle conduit à adopter divers formats qui visent à permettre l'expression de nuances. Ces formats, apparentés aux échelles de Likert¹, comportent un nombre variable de cases à cocher, constituant une structure ordinale entre deux polarités extrêmes.

Un des modèles les plus populaires, employé assez systématiquement dans les sondages d'opinion, est le suivant^a :

	Pas d'accord	Plutôt pas d'accord	Plutôt d'accord	D'accord
Êtes-vous d'accord avec telle opinion ?	☐	☐	☐	☐

On peut trouver aussi des échelles dont seules les extrémités sont verbalisées², c'est-à-dire ornées d'un texte explicite :

Souhaitez-vous telle chose ?	<i>beaucoup</i>	☐	☐	☐	☐	☐	<i>pas du tout</i>
------------------------------	-----------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------

On peut encore trouver des échelles où les extrémités et les positions intermédiaires sont verbalisées³ :

Les tribunaux sont trop sévères			Les tribunaux sont trop laxistes		
☐	☐	☐	☐	☐	☐
Tout à fait d'accord	Largement d'accord	Plutôt d'accord	Plutôt d'accord	Largement d'accord	Tout à fait d'accord

L'observation attentive de ces échelles révèle des différences notables : le nombre de cases peut-être pair ou impair ; les positions intermédiaires peuvent être verbalisées ou non. Le dernier cas présente la situation, potentiellement troublante, d'une échelle de verbalisation intermédiaire symétrique, la sémantique de la réponse étant décomposée entre la *polarité*, notée par les extrémités, et l'*intensité* notée par les positions.

La question de savoir si le nombre de cases doit être pair ou impair n'est peut-être pas seulement une querelle de *small-endians* et de *big-endians*⁴. La critique

1. Ensembles de catégories ordonnées, proposées comme réponse à un item bipolaire. Cette technique est employée par R. Likert dès 1932, spécifiquement dans la démarche de mesure d'attitudes.

a. NDLR : Le format de la revue oblige à couper les consignes ; cela est à éviter.

2. Exemple emprunté à Laveaux et Grégoire [1997]

3. Exemple adapté de Rouanet, Le Roux et Bert [1987]

4. Cf. *Gulliver's Travels*, de Jonathan Swift, où il est débattu s'il faut entamer l'œuf à la coque par le gros bout ou par le petit bout.

adressée aux échelles impaires est qu'elles autorisent une position centrale, largement exploitée par les indécis ou ceux que la question n'intéresse pas, et de ce fait sur-représentée dans les résultats. Les échelles paires apparaissent alors comme une nouvelle tentative (plus douce que dans le oui / non) de forcer l'enquêté à se déterminer. Le succès n'est pas garanti, car on voit des répondants résister à cette pression en plaçant une croix *entre* les deux cases centrales, signe probable d'une certaine mauvaise humeur et en tous cas d'un refus de la consigne.

Un problème plus fondamental est que, même si l'on code la position sur ces échelles par un nombre au moment du traitement, elles ne constituent pas pour autant des variables *métriques* justiciables de tous les (mauvais) traitements, y compris moyennes, écart-types et corrélations, mais seulement des variables *ordinales*, pour lesquelles l'appareil statistique est nettement moins fourni. Ceci est particulièrement sensible avec les échelles qui comportent plus de 4 positions : si celles-ci sont verbalisées, on est amené à faire usage d'adverbes modalisants dont la valeur sémantique n'est pas nécessairement la même pour tous les lecteurs, et dont la qualité même d'intermédiaire n'est pas évidente : dans la troisième échelle, dira-t-on que « large-ment d'accord » est exactement situé entre « d'accord » et « tout à fait d'accord », ou plutôt proche de l'un ou de l'autre ? Vraisemblablement cela dépend des répondants.

2.2.2 LE CLASSEMENT : UNE FAUSSE BONNE IDÉE

L'ingéniosité des chercheurs et des apprentis-chercheurs ne connaissant pas le repos, on rencontre souvent, devant la difficulté à obtenir des appréciations absolues, des tentatives pour recueillir des appréciations relatives, autrement dit des *préférences* ou des *classements*. Le recours à cette démarche est souvent justifié par le désir de varier les modalités de questionnement, voire d'égayer le questionnaire avec une pincée de jeu. Habituellement, ce genre de question présente une petite dizaine d'items, avec une consigne telle que « Numérotez par ordre d'importance (ou de préférence) décroissante les 3 (ou 4, ou 5...) éléments qui vous paraissent les plus importants ».

Ces questions, du type « classer p objets pris parmi n objets », qui paraissent méthodologiquement anodines, soulèvent en fait des cascades de problèmes techniques et épistémologiques. Une première objection, d'ordre ergonomique, consiste à faire remarquer que l'opération de classer p objets selon un critère se laisse décomposer en $p(p-1)/2$ opérations élémentaires de comparaison des objets deux à deux. C'est donc une opération dont le coût croît comme le carré du nombre d'objets à classer, avec d'éventuelles conséquences sur la fatigue du répondant.

Un second problème apparaît quant au codage des réponses en vue de leur exploitation statistique. En se plaçant dans l'hypothèse favorable où tous les enquêtés ou presque ont respecté la consigne, et consciencieusement classé exactement p items, sans *ex æquo* ni lacunes dans les rangs, l'enquêteur se trouve confronté à un choix

radical : ou bien il crée p variables nominales du style « item classé en premier », « item classé en second », etc., chacune de ces variables possédant n positions, à savoir une position (ou étiquette) différente pour chaque item, ou bien il crée n variables numériques, une pour chaque item possible, du style « rang de l'item a », « rang de l'item b », etc.

Dans le premier cas, la variable « item classé en premier » fournira effectivement un palmarès des items, du moins quant à leur apparition en première place. En revanche, il sera assez difficile de prendre en compte le fait qu'un item très souvent classé en seconde ou troisième place mérite peut-être une attention particulière. Conscients du problème au moment de l'exploitation (mais pas avant), beaucoup recoderont simplement les données pour compter combien de fois un item fait partie des p items classés. Cela donnera une base plus solide à l'interprétation, en abandonnant au passage l'information propre au classement : en bref, les enquêtés auront travaillé en partie pour rien.

Dans le second cas, le codage des rangs laisse espérer la possibilité de calculer un rang moyen par item, et donc des comparaisons entre items prenant en compte la totalité de l'information, à ceci près que le rang d'un objet qui n'a pas été classé n'est pas défini, et qu'il faudrait donc que chaque sujet classe tous les objets pour que l'opération se déroule sans souci.

De plus, les cas d'infraction à la consigne sont assez fréquents avec ce genre de questions, peut-être en raison même de leur coût en charge mentale. Certains répondants pressés ne remarquent pas qu'il faut classer, et mettent seulement des croix là où ils sont d'accord ; d'autres, scrupuleux, ne parviennent pas à choisir entre deux items et les classent *ex æquo* ; d'autres classent plus ou moins d'items que ce qui leur est demandé : parfois d'ailleurs la consigne omet de préciser combien d'items il faut classer, et en laisse le choix aux répondants. Ces infractions peuvent avoir des conséquences fatales sur l'exploitabilité de la question, qu'on peut être amené à abandonner purement et simplement, cas extrêmement caractéristique d'un gaspillage d'efforts qui aurait pu être évité. Même la position de repli consistant en un codage binaire classé/non-classé avec une variable logique par item n'est pas entièrement satisfaisante, puisque la statistique finale est biaisée par le poids plus important donné aux « gros classeurs », qui tendent à classer beaucoup d'items, qu'aux « petits classeurs », plus économes de leurs efforts de classement.

2.2.3 COMMENT S'EN TIRER ?

A ceux qui me consultent *après* s'être placés dans ce mauvais pas, je suggère parfois un expédient qui permet de récupérer quand même ces données, au prix d'une hypothèse un peu forte mais salvatrice. Cette hypothèse est que les objets non classés sont, dans le système de préférences exprimé, nécessairement placés *après* le dernier objet classé. De plus, comme aucune indication n'est donnée sur leurs

positions respectives, on peut supposer qu'ils sont situés ensemble au même rang virtuel. Comment déterminer ce rang ? Si on le place juste après le dernier objet explicitement classé, on ne fait pas assez de différence entre le fait de classer et le fait de ne pas le faire. Si on le place arbitrairement loin, très loin, par exemple au 20^{ème} rang quand il fallait classer 5 objets, on retrouvera les biais entre gros et petits classeurs. En revanche, si on détermine ce rang comme le rang moyen parmi les rangs non attribués, chaque répondant disposera du même pouvoir de classement, puisque la somme des rangs attribués sera une constante.

Supposons qu'un enquêté ait classé 3 objets parmi 10, quelle qu'ait été la consigne. Les trois objets classés auront les rangs 1, 2 et 3, et les objets non classés auront tous le même rang $(4 + 10) / 2$, soit le 7^{ème} rang⁵. Un autre enquêté a classé 5 objets, en leur donnant les rangs 1, 2, 3, 3 et 4, pour manifester son hésitation entre deux items. On rectifiera l'expression défectueuse de l'*ex æquo* en codant ces cinq objets aux rangs 1, 2, 3^{1/2}, 3^{1/2} et 5, et les cinq objets non classés au rang moyen $(6 + 10) / 2$, soit le 8^{ème}. Pour l'enquêté distrait qui a seulement coché 4 items sans les classer, on accordera à ces quatre items le rang moyen $(1 + 4) / 2$ soit 2^{1/2}, et le rang $(5 + 10) / 2$ soit 7^{1/2} au reste. Dans le cas particulièrement pervers où un enquêté a classé certains items et seulement coché certains autres, on peut raisonnablement considérer que cocher manifeste une préférence moindre que classer, et ranger les items cochés au rang moyen après les items classés, mais avant les items non-classés. On constatera que dans tous les cas le total des rangs attribué après rectification par chaque enquêté est toujours le même, soit 55 dans le cas de 10 objets, et plus généralement $n(n + 1) / 2$, qui est la somme des n premiers entiers. Même le cas où le répondant n'a rien classé peut théoriquement être transcrit en n objets uniformément placés au rang $(n + 1) / 2$. Ne s'agirait-il pas là d'une violence exorbitante que d'arracher une réponse à quelqu'un qui n'a pas répondu ? Cela peut se discuter. Classer tous les objets au même rang peut apparaître comme une manière correcte de coder l'indifférence. La violence n'a-t-elle pas commencé quand on a demandé aux enquêtés d'effectuer ce travail qui ne leur est d'aucun profit ? Peut-on distinguer, parmi ceux qui observent la consigne, ceux qui le font parce qu'ils acceptent la question de ceux qui ont simplement l'habitude d'appliquer une consigne parce qu'on leur demande ? Gardons ce débat de côté pour le moment où nous n'aurons plus d'autre problème.

Reste, au-delà des pirouettes arithmétiques, que le rang de classement d'un objet ne mesure pas une caractéristique intrinsèque de l'objet, mais seulement le nombre (augmenté de 1) d'objets qu'on lui a préférés, ce qui dépend évidemment de quel

5. Plus exactement $(4 + 5 + 6 + 7 + 8 + 9 + 10) / 7$, soit la somme des rangs restant à utiliser divisée par le nombre d'objets candidats. Il est facile de montrer que les deux modes de calcul sont équivalents parce que les rangs inutilisés constituent une progression arithmétique.

choix d'objets était proposé. En d'autres termes, les variables de classement, en dépit de leur codage numérique, sont foncièrement ordinales, et le caractère hautement suspect de leurs qualités métriques doit inciter à une certaine retenue dans l'exploitation mathématique qu'on en fait.

Pour fixer les idées sur ce point, si on me demande de classer par ordre de préférence décroissante quatre plats désormais traditionnels de notre cuisine nationale, à savoir le goulash, la paëlla, le couscous et les spaghetti *alla bolognese*, je répondrai sûrement, dans l'ordre, couscous, goulash, spaghetti et paëlla. Toutefois le fond de ma pensée est que j'aime vraiment beaucoup le couscous et que je suis capable d'en consommer sans modération, que j'apprécie un goulash et une sauce bolognaise s'il s'agit de recettes authentiques et non de succédanés industriels, et que je ne supporte absolument pas la paëlla parce que les fruits de mer me répugnent. D'ailleurs, entre le goulash et les spaghetti, je serais certainement capable, certains jours, d'hésiter longuement. Il s'ensuit que toute manipulation mathématique qui consisterait au fond à considérer que les écarts respectifs entre couscous et goulash, entre goulash et spaghetti, entre spaghetti et paëlla se valent tous, et constituent une sorte d'*échelle d'intervalles*, s'avérerait parfaitement abusive. Dans mon cas personnel, l'écart entre goulash et spaghetti est assez ténu, tandis que l'écart entre spaghetti et paëlla est un véritable gouffre, qui ne saurait en aucune façon se comparer à l'écart entre couscous et goulash.

2.3 DES ÉCHELLES ET DES PRÉFÉRENCES À LA VALUATION NUMÉRIQUE

En fait, si l'on veut obtenir une véritable mesure de mes préférences culinaires, *il suffit peut-être de me le demander*. Je suis parfaitement capable de situer ces plats sur une échelle d'évaluation graduée de 0 à 10, ou de 1 à 100 ou tout ce qu'on voudra. Si on veut m'obliger à peser finement mes choix, le mieux est de me donner 100 F à dépenser librement parmi le choix de quatre plats. Sans doute affecterai-je alors 50 F au couscous, 26 F au goulash, 24 F aux spaghetti et 0 F à la paëlla ⁶. En tous cas, si l'on souhaite un classement, ma réponse le contient, et si on veut comparer les différences entre mes choix, on peut légitimement le faire.

Ce qu'on me demande, c'est une *valuation* numérique de différents objets, c'est-à-dire l'association à chaque objet une valeur numérique ⁷. Difficile ? Nous le faisons tous sans cesse. Le modèle de la notation scolaire est si prégnant chez les enfants d'âge scolaire et les adolescents qu'ils l'utilisent spontanément pour évaluer les choses les plus diverses, y compris le charme personnel de leurs contemporain(e)s. Une enquête en milieu scolaire menée récemment sur les pratiques de jeux vidéo ⁸,

6. Peut-être même dans ce dernier cas serais-je prêt à payer pour ne pas devoir en manger.

7. On aura bien noté que je prends bien soin de ne pas parler d'évaluation, processus complexe, polysémique, largement étudié par des pléiades d'auteurs. Je ne m'intéresse ici, dans la *valuation*, qu'au résultat : le nombre qui value.

utilisant des consignes de valuation d'items par un nombre entre 0 et 10, n'a nécessité aucune explication verbale de la consigne. A côté de ces valuations en scores⁹, il est envisageable d'employer des valuations en pourcentages, fréquemment utilisés par les adultes comme mesure d'adhésion à une opinion, de probabilité d'un événement ou du degré de confiance à attribuer à une information: « Sur ce que vous venez de dire, je suis d'accord à 80% ».

Si l'on garde à l'esprit le modèle des 100 F à répartir entre différents achats, on peut être tenté de proposer au répondant de valuer sur la base 100 une série d'items, avec une sur-consigne précisant que le total pour la série doit faire 100. Des expériences menées avec les étudiants de Sciences de l'Éducation montrent que cette consigne est trop lourde à respecter, même chez des enquêtés instruits, puisqu'une part non négligeable de répondants offre des totaux différents de 100. Le souci de fixer un total constant à la valuation d'une série fermée d'items renvoie à la préoccupation déjà exprimée de ne pas surpondérer l'opinion des « gros répondants », mais se heurte à la désuétude du calcul mental chez nombre d'adultes, et la démarche doit être rejetée pour des raisons ergonomiques. Toutefois, il est aisé¹⁰ de transformer après coup les valuations sur une série par une règle de trois ramenant au total souhaité.

3 UNE PROPOSITION PRATIQUE

L'ensemble de ces considérations, et le souci de ne pas surcharger la tâche déjà difficile de l'apprenti chercheur par l'effort de choisir une technique parmi de nombreuses variantes, m'amène à proposer ici un instrument unique de recueil de données, qui, accompagné des techniques appropriées de traitement statistique, semble pouvoir répondre à la plupart des attentes énumérées ci-dessus.

3.1 UN VALUATEUR NOMMÉ V8

Ce qu'on nommera désormais un « valueur de type V8 » est la réunion d'une proposition textuelle, d'une grille horizontale de 8 cases et de verbalisations polarisées aux extrémités de ladite grille.

8. Dedourge, Dubus, Fichez & Pauwels, [1998].

9. Ce terme paraît plus approprié à une valuation sur une échelle de 0 à 20, comme il est de coutume d'employer à partir de la 6^{ème}. En effet, le mot « score », fréquemment utilisé dans le sens de « compte des points dans une compétition sportive », a très vraisemblablement le sens originel de *vingt*. Comme tel, il figure dans le mot anglais *threescore*, qui signifie soixante. Comme on ne retrouve pas d'équivalent dans les langues germaniques les plus proches (Allemand et Néerlandais), il s'agit sans doute du vestige d'une numération celtique en base 20, encore présente en français de France dans le nombre *quatre-vingt*, quand des nations plus marquées par l'esprit de système disent *octante*.

10. A condition de disposer d'un logiciel orienté dans le sens d'une saisie sans censure des données brutes, avec de grandes facilités de définition de nouvelles variables à partir des variables d'origine par des formules simples d'emploi. C'est une des ambitions du logiciel Adso, (Dubus [1998a]).

	complètement faux		tout à fait exact
Enseigner les statistiques en Sciences de l'Éducation est une entreprise insensée			

La consigne de réponse à une telle question est de tracer une croix (ou de cocher) dans la case qui correspond le mieux. Cette consigne (on y reviendra) peut être fournie une fois pour toutes dans un questionnaire.

Seules les extrémités du valuateur portent un commentaire, ou *verbalisation*. Pourquoi ? Parce que dans le cas contraire les positions intermédiaires seraient inévitablement garnies d'adverbes approximatifs (un peu, plutôt, nettement ...) dont la valeur n'est pas attestée comme constante auprès de différents locuteurs, mais aussi parce que cela prend, inutilement, beaucoup de place.

Le cœur du propos est d'obtenir une réponse *spatiale* à la tension du répondant entre les deux polarités de la réponse. Les traits entre les cases ne sont d'ailleurs pas indispensables du point de vue du répondant, mais ils sont bien utiles pour celui qui dépouille le questionnaire. C'est pour une raison similaire que le valuateur comporte huit cases. Huit est un compromis entre le souhait d'un nombre de cases assez important pour que ce nombre échappe à l'*insight computing*¹¹, ou à tout le moins exige un effort soutenu pour être converti en nombre, et la place que l'échelle prend dans la page avec des cases approximativement carrées et d'une hauteur suffisante pour éviter d'avoir à faire un effort pour viser. Pourquoi essayer d'éviter la conversion mentale en nombre, dans le même temps où on n'offre pas d'équivalent verbal pour chaque case ? Précisément pour que la réponse soit une *position*, le résultat d'un geste et pas d'une analyse¹².

Toujours pour les mêmes raisons, ce valuateur suppose que la polarité faible, ou négative, soit à gauche, et la polarité forte, ou positive, à droite, parce que c'est le sens de notre écriture. Cette orientation ne doit pas varier au cours du questionnaire. Les verbalisations polarisées ne sont là que pour fournir une sémantique vraisemblable au geste de cocher. Il serait extrêmement contre-productif d'attendre du ré-

-
11. Le nombre d'objets pour lesquels on est capable d'évaluer le cardinal de l'ensemble sans énumérer, d'un coup d'œil.. Les chiens savent faire ça pour trois objets : essayez dans un bistrot de ne donner que deux morceaux de sucre à l'épagneul qui vous regarde fixement avec la misère du monde dans son œil larmoyant ! Les êtres humains le font, paraît-il, assez facilement jusque vers 6 ou 7, plus difficilement ensuite, et encore pas tous.
 12. Sollicité de fournir une justification théorique à ce choix, je ne m'en sens pas actuellement capable, et je préfère dire provisoirement qu'il s'agit d'une intuition. Toutefois, il existe des convergences avec d'autres démarches : une émission d'information médicale montrait l'emploi devaluateurs de ce genre dans le cadre d'un dialogue d'un patient avec son médecin spécialisé dans la prise en charge de la douleur. Dans un ordre d'idées voisin, une enquête sur les représentations de l'avenir dans une population de chômeur utilisait en 1984 le tracé de *chemins* entre des maisons représentant des lieux et moments de la vie comme modalité de réponse non-verbale (Abécédaire [1985]).

pendant qu'il les lise avec attention, et que leur signification exige une inversion du sens de $-$ vers $+$, à moins évidemment que l'objet du questionnaire soit de mesurer l'attention du répondant.

Les cases sont contiguës, et non pas séparées, pour suggérer le continuum spatial. Si plusieurs valuateurs, munis des mêmes verbalisations, peuvent constituer une série, on les structure en un seul tableau, de la manière suivante :

	complètement faux				tout à fait exact			
Enseigner les statistiques en Sciences de l'Éducation est une entreprise insensée								
Une thèse sans statistique est comme un baiser sans moustache								
Les statistiques doivent être confiées aux mécaniciens-graisseurs spécialisés								
Avec les statistiques on peut prouver absolument n'importe quoi								

Le but de la manœuvre est de détourner l'attention de l'instrument de recueil, de le rendre tellement omniprésent qu'on ne le voit plus. On tâchera de trouver des propositions qui tiennent sur une ligne. Un saut de ligne introduirait une solution locale de continuité dans l'organisation spatiale susceptible de déconcentrer le répondant. L'idéal est que celui-ci, une fois familiarisé avec le procédé, se concentre sur le texte en une ou deux fixations oculaires, et rende sa réponse sous la forme du geste de cocher dans la grille, sans même avoir besoin de la regarder.

3.2 PERSPECTIVES D'EXPLOITATION

Quel genre de donnée produit unvaluateur? Formellement, un nombre, de 0 à 7 ou de 1 à 8. Généralement entier, mais si un petit malin d'enquête place une croix à cheval sur deux cases, pourquoi ne pas coder en $1/2$ points? S'agit-il pour autant d'une mesure au plein sens du terme? Probablement plus que dans des échelles verbalisées à chaque degré. Pourtant, on verra ci-dessous que la prudence commande de n'exploiter in fine que la qualité ordinale des données recueillies. Des traitements statistiques appropriés vont permettre de contourner quelques-uns des pièges classiques avec ce genre de données.

Une des premières difficultés, qui ne tient pas à notre technique de recueil de données mais à laquelle celle-ci n'échappe pas, est la question de la désirabilité sociale de la réponse, avec sa variante, le « politiquement correct ¹³ » du milieu enquêté.

13. Dans les années 70 et 80, j'ai eu à participer avec mon ami Jacques Hédoux au recrutement académique

Imaginons qu'on ait posé, dans une enquête en milieu enseignant, la question suivante :

complètement faux tout à fait exact

Tout ce que nous faisons est fait dans l'intérêt des élèves
☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

On obtiendrait vraisemblablement une distribution des réponses de ce genre :

0	1	2	3	4	5	6	7
0 %	0 %	0 %	0 %	3 %	9 %	28 %	60 %

Qu'en déduire ? Si l'on s'en tient aux modèles classiques, ceux des sondages d'opinion, on coupe au milieu (entre 3 et 4), et l'on conclut que la totalité (100 %) des enseignants enquêtés tend à adhérer à la proposition « Tout ce que nous faisons est fait dans l'intérêt des élèves ». Beau scoop ! Ici la proposition à valuer est tellement évidente que la seule opposition pertinente est entre la position 7 et n'importe quelle autre position : toute nuance par rapport à l'adhésion inconditionnelle peut légitimement être interprétée comme un refus plus ou moins affirmé de la proposition (et peut-être du quantificateur « Tout » au début de la proposition). Le commentaire valide devient : 60 % des enseignants enquêtés disent adhérer *sans réserve* à la proposition ... Inversement, la question

complètement faux tout à fait exact

Notre ministre actuel manifeste un grand respect pour les enseignants
☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

des Conseillers en Formation Continue, enseignants du second degré, ou plus rarement du premier, qui s'ils étaient recrutés quitteraient leur(s) classe(s) pour être détachés dans des fonctions d'animation, d'organisation et de négociation de stages de formation professionnelle continue dans les établissements publics. La circulaire organisant le recrutement stipulait : « au cours de l'entretien, on s'assurera que le candidat se rapproche de la formation continue par une démarche positive vis à vis de la formation initiale ». En clair, il s'agissait de démasquer les déserteurs, c'est-à-dire les enseignants qui en ont assez d'enseigner. Évidemment, tous les candidats avaient épluché la circulaire ; aussi, nombreux étaient ceux qui au cours de l'entretien de recrutement vantaient l'intérêt des *retombées pédagogiques* des expériences qu'ils avaient pu avoir en formation d'adultes *sur la formation initiale*. Un discours moins contraint aurait pu être : « Ecoutez ! J'enseigne depuis 10 ou 15 ans en collège ou en lycée professionnel. J'ai fait ce métier du mieux que j'ai pu, mais maintenant, j'aimerais m'ouvrir à d'autres genres de tâches et de responsabilités. Il paraît que vous avez un job intéressant à proposer. De quoi s'agit-il ? ». Dans le moment même où ils entreprenaient ostensiblement une démarche *devant* les conduire à quitter la fonction d'enseignant, ces candidats ne pouvaient tout simplement pas dire qu'ils *souhaitaient* la quitter. (Dubus [1980]).

aurait pu, à certaines époques, recueillir une distribution des réponses inverse. Dans ce cas, toute réponse autre que 0 pourrait être interprétée comme une timide approbation de l'œuvre du susdit. Que retirer de ces deux exemples ? Que l'interprétation des résultats d'un valuateur peut différer d'une question à l'autre, et qu'elle dépend de l'allure de la distribution brute enregistrée sur l'ensemble des répondants. On retrouvera plus loin cette préoccupation à propos des dichotomies. Si on veut continuer à étayer le commentaire par référence à la donnée brute et éventuellement à la coupure du milieu, alors il faut rédiger des propositions qui tendent elles-mêmes à partager l'opinion dans des proportions sinon égales, du moins comparables. Le problème est que par définition on ne sait pas avant d'avoir enquêté comment se répartiront l'opinion, les représentations ou les attitudes. Il est donc sans doute plus intéressant de pouvoir rectifier *a posteriori* ces effets d'attitude dominante.

Un autre facteur de distorsion est la possibilité que certains individus adoptent un style propre d'usage du valuateur, comportant par exemple une tendance à valuer plutôt vers la droite, ou plutôt vers la gauche, à utiliser largement toute l'échelle, ou au contraire à concentrer ses valeurs vers le centre ¹⁴. Une valeur 7 chez un individu usuellement généreux en cotation ne mérite pas la même attention qu'une valeur 7 chez un individu pour qui elle est exceptionnelle. Là aussi, il peut s'avérer intéressant de neutraliser les variations individuelles.

On évolue ainsi vers un modèle où la réponse brute d'un individu i à un valuateur j se laisse décomposer en trois éléments :

$$b_{ij} = s_i + d_j + p_{ij}$$

où b_{ij} est la valeur brute, telle qu'elle a été saisie, s_i la composante de style personnel de l'individu i , d_j la composante liée à l'opinion dominante concernant l'item j dans l'échantillon d'individus étudié, et p_{ij} la composante propre traduisant la réaction spécifique de l'individu i devant l'item j .

Une méthode brutale consisterait à calculer comme estimateur de s_i la moyenne des valeurs utilisées par l'individu i sur l'ensemble des valuateurs présents dans le questionnaire ¹⁵, puis à calculer comme estimateur de d_j la moyenne des valeurs utilisées par l'ensemble des répondants sur le valuateur j , et enfin à calculer chaque p_{ij} par soustraction. Cependant le caractère initialement ordinal des valeurs brutes nous incite à ne pas fonder trop d'espoirs sur la pseudo-métrique engendrée par nos mani-

14. On trouve là quelque chose de similaire au « tempérament docimologique » d'un correcteur.

15. Cela suppose bien sûr que le valuateur soit utilisé un nombre assez important de fois dans le questionnaire. On peut aussi n'utiliser que la moyenne des valeurs sur un sous-ensemble des valuateurs, si on a des raisons de penser qu'à cause des thèmes particuliers abordés dans ce sous-ensemble l'attitude spécifique des répondants peut varier par rapport à d'autres parties du questionnaire, et varier de manière différente selon les individus. De nombreuses variantes, plus sophistiquées les unes que les autres, sont possibles, aussi longtemps qu'on sait encore de quoi on parle.

pulations. Aussi paraît-il plus approprié de procéder en deux temps : dans un premier temps on neutralise l'effet du style individuel en calculant pour chaque individu et pour chaque item valué la différence entre la valeur brute et la valeur moyenne de l'individu pour l'ensemble des items. Cette première étape peut être omise si l'étude de la distribution moyenne sur les items montre peu de différences interindividuelles. Dans un second temps, sur ces différences (ou sur les valeurs brutes si on a omis la première étape), on calcule des variables logiques qui valent VRAI si la valeur étudiée est supérieure à la *médiane* de la distribution des valeurs pour l'ensemble des individus, et FAUX sinon. Les quelques exemples qui suivent (pour lesquels on travaille sur des valeurs brutes) montrent que la coupure dichotomique peut intervenir en une quelconque position inter-valeurs :

Avec une distribution d'origine très asymétrique

Scores d'origine	0	1	2	3	4	5	6	7
Distribution des réponses	0%	0%	0%	0%	3%	14%	34%	49%
Valeurs dichotomiques	FAUX							VRAI
Distribution des réponses	51%							49%

Avec une distribution d'origine à peu près « plate »

Scores d'origine	0	1	2	3	4	5	6	7
Distribution des réponses	18%	7%	11%	12%	15%	9%	20%	8%
Valeurs dichotomiques	FAUX				VRAI			
Distribution des réponses	48%				52%			

La transformation en valeurs logiques conserve le caractère ordinal des données initiales, et renonce à toute illusion sur leurs qualités métriques. L'intérêt d'opérer la dichotomie sur la médiane ¹⁶, et non sur la moyenne, est qu'elle fournit par définition le découpage de la population le plus proche possible de l'équipartition, même avec des distributions fortement asymétriques. Or la théorie de l'information nous dit que le critère qui contient le plus d'information est celui qui partage la population en deux groupes d'effectifs égaux. Concrètement, l'espérance de voir se dégager des tendances et des relations lors de croisements avec d'autres variables est toujours bien meilleure avec des contrastes équilibrés en effectifs.

Le résultat de ces dichotomies sur la médiane est un système de variables dont l'interprétation s'effectue à un degré d'abstraction évidemment supérieur à celui de

16. Le caractère un peu technique de cet exposé ne doit pas effrayer le lecteur : ces opérations de transformations de variables sont rendues à peu près automatiques par les logiciels appropriés comme Adso (Dubus [1998a]).

la donnée brute. Cela peut se formaliser ainsi : « compte tenu de son style personnel de réponse et de la tendance dominante dans la population, cet individu a plutôt tendance à conférer une valeur élevée à cet item ». À condition de ne pas oublier dans le commentaire à quel degré d'abstraction on se situe, plusieurs traitements intéressants peuvent ensuite être effectués sur ces variables. Le propos principal étant ici un point de méthode relatif au recueil de données, on sera volontairement peu disert sur ces traitements, qui justifieraient un vaste développement abondamment muni d'exemples. Ces traitements se rattachent d'une manière ou d'une autre à deux types de démarche : la comparaison entre items et la comparaison entre sujets.

Les items obtenus à ce stade des opérations sont représentés par des variables logiques. Les variables logiques, qui rassemblent les qualités des variables nominales (avec les positions Vrai / Faux), des variables ordinales (deux valeurs sont nécessairement ordonnées) et des variables métriques (0/1) autorisent la plus large gamme de traitements et de tests (χ^2 de contingence, γ d'ordonnement, ρ et r de corrélation ...) ainsi que les mesures qui leurs sont propres, comme celles qui relèvent du modèle de l'implication¹⁷. La comparaison d'items, par exemple au moyen de tests de corrélation et d'analyses en composantes principales, peut permettre d'identifier des ensembles cohérents d'attitudes ou d'opinions.

Symétriquement, la comparaison entre sujets définis par leur profil de valeurs sur une série de variables dichotomiques issues de valuateurs peut déboucher sur des typologies de sujets, autrement dit l'identification de groupes disjoints de sujets aussi semblables que possible à l'intérieur des groupes et aussi dissemblables que possible de groupe à groupe. Les variables créées par typologie, qui sont des variables nominales, peuvent à leur tour être croisées avec d'autres variables issues d'autres parties du questionnaire.

3.3 MISE EN ŒUVRE CONCRÈTE ET PRÉCAUTIONS

3.3.1 PASSER LA CONSIGNE

Aussi simple d'emploi qu'elle soit, la technique des valuateurs réclame un minimum de soins et de précautions pour apporter pleine satisfaction. Un point très important est l'explication de la consigne de réponse. Comme un questionnaire comportant des valuateurs en compte souvent un nombre assez important, il n'est pas spécialement coûteux de ménager un espace non négligeable pour la présentation formelle de la consigne. Éric Liedts, dans l'excellent questionnaire de son *Mémoire de Maîtrise*, y procède avec un exemple de proposition à valuer, légèrement humoristique :

Pour répondre à ce genre de question, placer une croix dans la grille, à l'emplacement qui vous paraît le mieux exprimer le nuance de votre réponse. Par exemple :

17. Voir à cet égard Lahanier [1999] et Gras [1996].

complètement faux tout à fait exact

J'ai choisi cette unité parce que la salle où elle a lieu est la mieux chauffée

		X				
--	--	---	--	--	--	--

(ce n'est pas vraiment mon cas, mais pas tout à fait faux non plus)

On remarque qu'il place d'avance une croix dans le valuateur en position 2 (sur une échelle implicite 0–7), en la commentant par une verbalisation compliquée à dessein. Ce commentaire évite de transformer chaque valuateur en un équivalent verbalisé à chaque degré (conserver autant que possible le caractère spatial des réponses), et vise seulement à indiquer au répondant, de manière plutôt ludique, comment cette technique de réponse lui permet d'exprimer des nuances. Pour éviter la confusion avec les vrais valuateurs, le caractère utilisé pour le texte de la proposition est en italique, et l'ensemble de la consigne elle-même est placé dans un cartouche légèrement grisé, nettement séparé des séries de vrais valuateurs. A l'usage, l'effet d'explication paraît suffisant pour qu'il soit inutile de répéter la consigne dans le reste du questionnaire (4 pages A4 composées assez serré, présentées en A3 recto-verso plié par le milieu).

3.3.2 COMBIEN DE VALUATEURS ?

Le plein bénéfice de la technique suppose qu'on emploie des séries de valuateurs, structurées en un seul tableau. Quelle taille pour ces séries ? Il est clair qu'il s'agit de rechercher un compromis entre l'efficacité liée à l'habituation du répondant à la consigne et une possibilité de lassitude engendrée par la monotonie. Des séries de 10 propositions à valuer apparaissent à l'expérience comme un maximum qu'il vaut mieux ne pas dépasser. Un des facteurs d'acceptabilité pour une série de 6 ou 8 propositions à valuer est la cohérence sémantique des propositions de la même série, qui fait passer l'attention du répondant au niveau de la série perçue comme un ensemble et non plus du valuateur pris comme élément isolé. Il se peut que l'acceptabilité globale de la technique dépende de caractéristiques culturelles du public-cible ; toutefois les exemples de difficultés rencontrées sont contradictoires sur ce point, et on ne peut exclure que les échecs soient à relier au manque d'enthousiasme de l'apprenti-enquêteur pour la technique elle-même.

3.3.3 LA BEAUTÉ DU TEXTE

Reste que l'agrément ou le désagrément de la technique des valuateurs sont liés à d'autres facteurs que sémantiques ou idéologiques. La taille des cases en largeur et hauteur, la lisibilité de la fonte et du corps de caractère utilisé pour les propositions jouent certainement un rôle. Ainsi, la hauteur du corps de caractère relativement à la hauteur des cases doit être telle que soit perceptible l'unité physique de la série de valuateurs, les lignes de texte étant suffisamment séparées pour prévenir les erreurs de parallaxe, qui conduiraient par exemple à placer une croix sur la mauvaise ligne.

Ce type de recommandations, qui surprend toujours un peu les étudiants, pour qui un questionnaire est nécessairement un document de travail dont l'austérité traduit le sérieux de l'intention, procède de l'axiome selon lequel il n'y a pas de raison d'incommoder les gens, à qui on demande déjà l'effort de répondre, par une mise en page désagréable et triste, et que le questionnaire, instrument de communication écrite, doit d'abord être un beau document.

Instrument de communication, le questionnaire à valuateurs n'échappe pas non plus aux règles habituelles de clarté de la langue, d'absence d'ambiguïté des formulations, etc. Cette exigence est renforcée par le caractère unidimensionnel de l'instrument proposé au répondant pour exprimer des nuances : il convient que la proposition à valuer soit elle aussi unidimensionnelle par construction, qu'elle contienne une seule idée, dans une tournure affirmative et dépourvue de modalités telles que l'interrogation, la négation, la restriction, et, pire que tout, la combinaison de modalités.

3.3.4 LA NON-RÉPONSE

En dépit de toutes les précautions prises pour l'amadouer, il arrive que le répondant ne réponde pas. Que faire au moment du dépouillement ? Diverses solutions ont été proposées, toutes sujettes à caution. Coder une absence de réponse comme une position centrale dans le valuateur est à rejeter, compte-tenu notamment des transformations qu'on fera ensuite subir aux données, et qui pourront faire qu'une position initialement centrale ne le sera plus du tout ensuite. Une pratique apparemment plus ingénieuse consiste à reconstituer *a posteriori* par le calcul ce qu'aurait pu être la réponse qui n'a pas été donnée, compte-tenu des réponses données aux autres questions. Cette reconstitution d'un codage fantôme repose en fait sur un raisonnement circulaire : elle suppose que les relations entre les différentes réponses obéissent à une sorte ou une autre de régularité, ce que l'analyse a précisément pour objet de démontrer ; il s'agit donc d'un simple effet Münchhausen¹⁸. Il faut accepter l'idée que l'information qui n'a pas été fournie n'est pas disponible. On recommande parfois de prévoir une variable séparée codant le cas de non-réponse, manifestée ou non par une case à cocher (mais que faire du répondant qui ne coche même pas la case non-réponse ?). L'expérience montre qu'on n'en fait généralement pas grand-chose dans le traitement, ne serait-ce que parce qu'il est impossible d'interpréter la non-réponse, qui peut être due à l'incompréhension, au refus de répondre ou à la simple inattention. Seuls quelques cas très particuliers où l'un des objets de l'étude serait la réticence différentielle de certaines fractions de la population enquêtée à répondre à certaines questions pourrait justifier de se donner le mal de coder séparément la non-réponse¹⁹. Encore dans ce genre de cas reste-t-il un doute relatif à l'existence

18. Le Baron de Münchhausen, enlisé avec son cheval dans un marécage, sauve sa vie et celle de sa monture en se soulevant lui-même par les cheveux.

de répondants chez qui la docilité vis à vis de la consigne l'a emporté sur la répugnance à répondre. La seule attitude raisonnable est généralement de ne pas coder, c'est-à-dire de coder par une absence de valeur, qui aura pour effet que le sujet qui n'a pas répondu sera exclu de tous les calculs impliquant la variable en question.

3.4 PERSPECTIVES AVANCÉES

Une extension intéressante de la technique des valuateurs consiste à faire valuer par les répondants des propositions issues de discours qu'ils pourraient avoir tenus eux-mêmes, avec leurs propres mots. Pour cela, il convient de recueillir préalablement du matériau verbal par entretiens au sein de la population étudiée, puis d'en extraire par l'analyse de contenu des propositions, stabilisées en forme canonique à partir d'expressions jugées équivalentes, et autant que possible nettement polarisées vis à vis des thèmes liés à la problématique de l'enquête. Cette forme de coopération de l'entretien et du questionnaire, qui vise à amener une population à un retour réflexif sur son propre discours, peut déboucher sur des aspects extrêmement intéressants. Ainsi il est possible de proposer à la valuation dans le questionnaire plusieurs propositions issues de la même formule canonique, reprise telle quelle ou transformée en son contraire supposé. Au traitement, la force des corrélations positives ou négatives entre items va révéler la qualité des jugements du chercheur quant à l'équivalence ou à l'antagonisme des propositions, et donc sa compréhension du système des idées dans l'ensemble du discours étudié. Pour prendre un exemple simple, si l'adhésion à une proposition n'est pas corrélée négativement et fortement à l'adhésion à la proposition que le chercheur croit être son contraire, alors le chercheur s'est trompé dans son appréciation du système d'idées, et il vaudrait mieux qu'il n'exploite pas les items concernés. C'est un des rares cas où le chercheur possède une preuve au moins négative que les enquêtés ont compris sa question et qu'il a compris leurs réponses. Pour que cette preuve soit crédible, il faut bien sûr que les items supposés corrélés soient dispersés à travers le questionnaire dans un nombre assez important d'items. Dans ma thèse (Dubus [1980]), j'avais ainsi construit 120 items à valuer (en quatre positions, à l'époque), correspondant à 30 thèmes sous-jacents, avec pour chaque thème deux items sémantiquement extrêmes et deux items intermédiaires (du moins le croyais-je), les items étant ensuite disposés aléatoirement, avec une contrainte de non-proximité des items issus du même thème (au moins 5 items de distance). Le résultat de la validation par les corrélations s'avérait a priori assez cruel, puisque sur les 30 thèmes employés, 13 ne satisfaisaient pas au modèle,

-
19. Par exemple une étude sur le caractère plus ou moins tabou de questions relatives à la religion, à la politique, à la sexualité, à l'argent. A vrai dire dans ce cas les non-réponses n'en sont plus : l'enquêteur, dans une posture de ruse vis à vis de l'enquêté, feint de s'intéresser à la réponse de celui-ci, alors qu'il ne s'intéresse réellement qu'au fait qu'il réponde ou non. Il devient impossible de non-répondre, et les silences deviennent éloquentes.

rendu à vrai dire encore plus exigeant que la corrélation négative des extrêmes par la prise en compte des items intermédiaires. Ce résultat, d'abord perçu comme décevant, s'est avéré à la réflexion très satisfaisant, puisque les 17 thèmes qui avaient traversé victorieusement l'épreuve de la validation interne en sortaient évidemment très renforcés d'un point de vue argumentaire, et largement suffisants pour étayer une analyse consolidée des représentations dans la population étudiée.

4 POUR CONCLURE

Sans être le remède miracle à tous les problèmes de recueil de représentations dans un questionnaire, la technique du valuateur semble pouvoir remplacer avantageusement un certain nombre de bricolages sympathiques mais peu efficaces, tout en autorisant une prise en main *à peu près satisfaisante* dans un délai *raisonnable*. On mesure ici la relative modestie des objectifs, et le côté assez clairement prescriptif de la démarche : tout ce qui est recommandé ci-dessus n'est pas justifié dans le détail, et le lecteur est invité dans une large mesure à faire confiance à l'auteur. C'est que ceci s'adresse à des apprentis – chercheurs : on est effet ici dans l'ébauche d'une perspective de « pédagogie universitaire » appliquée à la méthodologie, et on mesure déjà ce qu'une telle démarche, au delà de la simple statistique, mobilise implicitement ou explicitement de réflexion sociologique, linguistique et ergonomique, bref de *décloisonnement*, comme on dit à l'école primaire.

Tout le mal que je souhaite enfin aux futurs compagnons-chercheurs, c'est d'être en mesure d'inventer ou d'adapter leurs propres outils de recueil de données et les méthodes de traitement qui leur conviennent, dès que possible et dans la plus grande autonomie possible.

BIBLIOGRAPHIE

- COLLECTIF ABÉCÉDAIRE (BAICHERE, CUKROWITZ, DUBUS, DUPREZ, RAHMANIA), 1985, « Calculs sioux pour réponses rusées », Lille, *Cahiers Lillois d'Économie et de Sociologie* n° 6.
- DEDOURGE C., DUBUS, A., FICHEZ, E., PAUWELS, G., 1998, *Potentialités éducatives des jeux vidéo à l'heure du multimédia*, Lille III, Rapport pour le Contrat de Plan État–Région.
- DUBUS A., 1980, *Les conseillers en formation continue sont-ils des agents de changement ?*, Université de Lille III, Thèse pour le Doctorat de 3^{ème} cycle en Sciences de l'Éducation.
- DUBUS A., 1998a, *Adso 3, logiciel et mode d'emploi*, Lille, Presses Universitaires du Septentrion.
- DUBUS A., 1998b, *Méthodes et pratiques du traitement statistique en Sciences Humaines. Étude de cas avec Adso 3*, Lille, Presses Universitaires du Septentrion.
- GRAS, R. et coll., 1996, *L'implication statistique, nouvelle méthode exploratoire de données*, Grenoble, Éditions La Pensée Sauvage.

- GHIGLIONE R. & MATALON B., 1991, *Les enquêtes sociologiques, Théories et pratiques*, Paris, Armand Colin.
- HOWELL D., trad. Fr. 1998, *Méthodes Statistiques en Sciences Humaines*, Paris, Bruxelles, De Bœck Université.
- LAHANIER D., 1999, *Conceptions du hasard et enseignement des probabilités et statistiques* Paris, PUF
- LAVEAUX D. & GRÉGOIRE J. 1997, *Introduction aux théories des tests en Sciences Humaines*, Paris, Bruxelles, De Bœck Université.
- LIEDTS É., 1998, *La licence des Sciences de l'Éducation : Étude de quatre actions sociales endogènes : Typologie des « étudiants-acteurs »*, Mémoire pour la Maîtrise de Sciences de l'Éducation, Université de Lille III.
- ROUANET H., LE ROUX B. & BERT M.-C., 1987, *Statistique en Sciences Humaines : Procédures naturelles*, Paris, Dunod.

